PROPOSAL FOR TOPIC APPROVAL TO BE SUBMITTED TO THE UNIVERSITY OF MUMBAI FOR THE DEGREE OF **Ph.D. (SCIENCE AND TECHNOLOGY)** IN THE SUBJECT OF **COMPUTER ENGINEERING**

Title of Topic	:	Development of Artificial Intelligence Framework for Academic Performance Forecasting
Name of the Student	:	Mr. Vijay Tukaram Patil
Basic Qualification	:	M.E.(Computer Engineering) B.E.(Computer Engineering)
In the Case of Ph.D. student Year of Passing of PET/NET/SET/GATE (Teaching Experience)	:	PET - 2024
Name of the Research Supervisor/ Guide	:	Dr. Bhavesh Valji Patel
Date of Superannuation / Retirement	:	20/03/2036
Name of the Research Center	:	Shah and Anchor Kutchhi Engineering College
Signature of Student	:	Patil
Signatue of the Research Supervisor	:	
Date of Proposal Submission		

Shah & Anchor Kutchhi Engg. College ★ Autonomous Institute of Mumbai

Head Director Principal
Research Centre

PRINCIPAL
Shah & Anchor Kutchhi Engineering College
Autonomous Institute
Mahavir Education Trust Chowk,
W. T. Patil Marg, Next to D-Mart,
Chembur - 400088

Sr. No.	Content	Page No.
1	Aim of Study and Objective of Study	1
2	Literature Review	2
3	National and International Status	9
4	Research Hypothesis	10
5	Research Methodology	11
6	Statistical Plan	13
7	Limitations of Study and Expected Outcomes	16
8	References	17

Statement by the Candidate

I wish to state that the work embodied in this synopsis titled "Development of Artificial Intelligence Framework for Academic Performance Forecasting" forms my own contribution to the work carried out under the guidance of Dr. Bhavesh Valji Patel at the Shah and Anchor Kutchhi Engineering College. I declare that this written submission represents my ideas in my own words and where others ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission.

(Candidate Signature)

Name: Mr. Vijay Tukaram Patil

Roll Number: 124PHCM1005

Aim of Study

Accurately forecasting Learner's performance is essential for enabling timely interventions and delivering personalized support. This study proposes a comprehensive approach that leverages both traditional Machine Learning algorithms and advanced Deep Learning models to forecast Learner academic outcomes [1]. Predictive analytics is utilized to estimate Learner's performance trends, identify potential dropouts, and analyze learning patterns. Prescriptive analytics provides strategic recommendations to enhance teaching effectiveness and optimize Learner's support mechanisms. The primary objective is to determine the most effective algorithms for predicting learner's academic performance and evaluating their impact on overall success.

Objective of Study

- To Identify and analyze key determinants- Academic, Demographic, and Behavioral—that significantly influence Learner performance, and to preprocess and enhance the dataset for effective Machine Learning modeling.
- To Investigate and assess Machine Learning algorithms utilized for evaluating Learner academic outcomes.
- To Implement and compare efficacy of various Machine Learning algorithms using Learner's data with features and find the most accurate one to help improve learning and learner's support.
- To Design a unified predictive framework that integrates Machine Learning and Deep Learning techniques for accurately forecasting Learner's academic performance.

Literature Review

Machine Learning (ML) and Deep Learning (DL) techniques have been extensively applied in predicting Learner's performance, addressing various challenges in educational Data Mining [7]. Machine Learning methods like Support Vector Machines and Decision Trees are often used to analyse Learner's academic and behaviour data. This helps schools and colleges build models to identify Learner who may need early help and support. Deep Learning methods, especially CNN, are good at finding complex patterns in large amounts of Learner's data. They have been used to study how Learner's interact with courses, giving better insights into their performance and learning habits. Machine Learning (ML) is a branch of Artificial Intelligence (AI) that enables systems to automatically learn and improve from experience without being explicitly programmed. It focuses on developing computer programs that can access data and use it to learn and derive meaningful patterns or functions. Machine Learning algorithms are widely applied in predicting educational outcomes and analyzing academic data.

Objective of the study:[1]

- To find out the gaps in existing predictive methods.
- To find out the attributes in analysing Learner's performance.

Machine Learning also has applications in the education sector. Many researchers use Machine Learning algorithms to evaluate and analyze Learner's academic performance using available educational data. These models are used for different purposes, such as predicting academic performance by classifying Learner's into low/medium/high, pass/fail, poor/good/best, etc. , predicting final marks , analyzing learning behavior, and early prediction of drop-out and retention rates , supporting Teachers and Institutions and monitoring the performance of Learner's. Machine Learning is classified into two main types: Supervised and Un-Supervised Learning, similar to how humans learn through guided (Supervised) or independent (Un-Supervised) methods.

The dataset gives a complete view of each Learner, helping to build better Machine Learning and deep learning models for predicting performance and understanding Learner's progress more clearly. The fundamental aim of Machine Learning is to enable machines to discern patterns, extract valuable insights, and enhance performance through accumulated experience. The Machine Learning process usually consists of three main components: Data, Model and Learning Algorithm.

- **Data:** The fundamental component of Machine Learning is Data. The process of training algorithms involves using big datasets that include instances, characteristics, or patterns that are pertinent to the job at hand. In order for Machine Learning models to be successful, both the quality and amount of the data present are crucial factors.
- **Model:** A model is a mathematical representation or structure that learns patterns from data. These models can be simple or complex depending on the nature of the work. Linear Regression, Decision Trees, Support Vector Machines, and Neural Networks stand as typical examples of prevalent models employed in Machine Learning.
- **Learning Algorithms:** Learning algorithms are the powerhouse of Machine Learning. Adjust model parameters based on input data to minimize errors and discrepancies between predicted outputs and actual results. Learning algorithms may be broken down into two primary categories: Supervised Learning, in which the algorithm is trained on data that has been labeled, and Un-Supervised Learning, in which the program detects patterns without the output of labels.

Machine Learning is categorized into Supervised and Un-Supervised Learning, similar to how humans learn with or without guidance.

Types of Machine Learning:

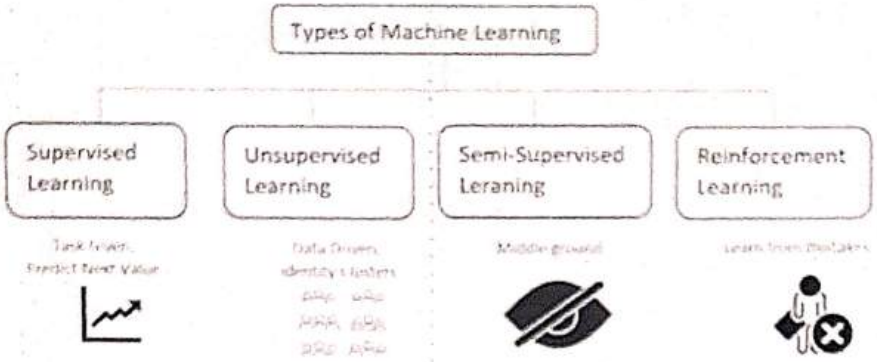

Fig 1: Types of Machine Learning

- **Supervised Learning:** Supervised Learning algorithms are Machine Learning techniques that rely on supervision, meaning the machine uses labeled data during the training process to learn and predict outcomes. Labeled data indicates that both the input and its corresponding output are known during training, enabling the model to make accurate predictions on new data. A well-trained model should produce outputs that closely match real-world results when presented with unseen data. For example, if a machine is trained on labeled images of bicycles and cars, it learns to recognize the distinguishing features of each. After training, when given a new image of a car, the model should correctly identify and classify it as a car if the training process was effective.
  Supervised Machine Learning algorithms have two main categories: Classification and Regression.
  Classification:
  - Predicts the label of a class
  - Predict the dataset's categories
  - Example: "Yes" or "No"

  Commonly Used Algorithms: Decision Tree Algorithm, Linear Regression, Random Forest Algorithm, Support Vector Machine Algorithm.

  Regression: Regression is used when the desired output is a continuous label. It means that the answer to the question is denoted by a quantity. These systems answer the questions of How much 'or How many'? for example what is the market price of this house¿ It can be determined based on the inputs of the model rather than being restricted to a set of possible labels.
  Commonly used algorithms: Decision Tree Algorithm, Simple Linear Regression Algorithm

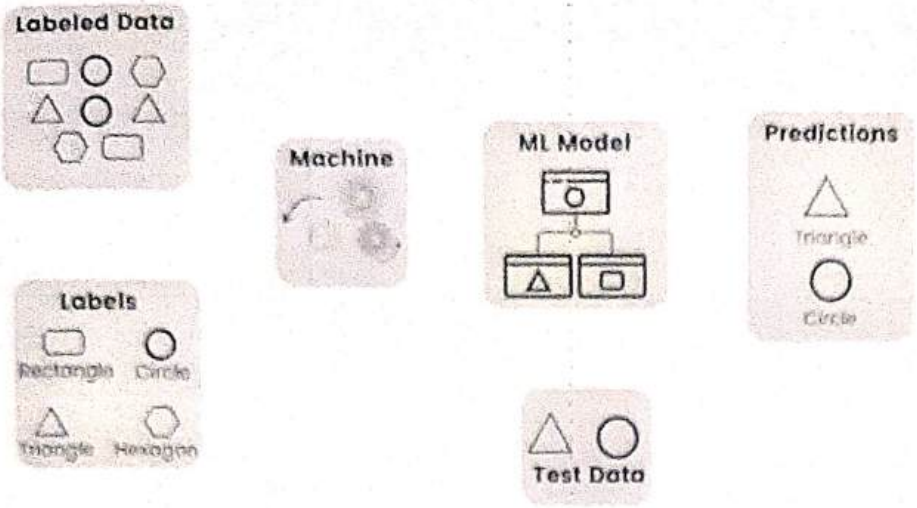

Fig 2: Supervised Learning (source: projectpro.io)

Types of Supervised Learning:

- **Support Vector Machine (SVM):** SVM is a supervised learning algorithm that finds the optimal boundary to separate classes, handling both linear and non-linear data effectively.
- **Logistic Regression:** Predicts the probability of a binary outcome (e.g., yes/no, true/false).
- **Decision Trees:** Creates a tree-like structure to classify data based on a series of decisions.
- **Random Forests:** An ensemble method that combines multiple decision trees to improve accuracy and reduce overfitting.
- **Neural Networks:** Complex algorithms that can learn intricate patterns in data, suitable for both classification and regression tasks.
- **K-Nearest Neighbors (KNN):** Classifies data points based on the majority class of their nearest neighbors.
- **Naive Bayes:** A probabilistic classifier based on Bayes' theorem with a simplifying assumption of feature independence.
- **Linear Discriminant Analysis (LDA):** Finds a linear combination of features that best separates different classes.

Supervised Machine Learning Applications:

- Fraud Detection
- Image Segmentation
- Medical Diagnosis
- Spam Detection

- **Un-Supervised Learning:** Unlike Supervised Learning, Un-Supervised Machine Learning does not require labeled data or supervision during training. Instead, it uses unlabeled data—data without predefined categories or outputs—allowing the model to independently discover patterns

and structures within the dataset. In unsupervised learning, the machine analyzes the input data and learns to classify or group similar items on its own. For example, if a machine is provided with various uncategorized images of toys, it will examine the features of each toy and group similar ones together without prior knowledge. Later, when given a new toy image, the machine should be able to assign it to the appropriate category based on the patterns it has learned.

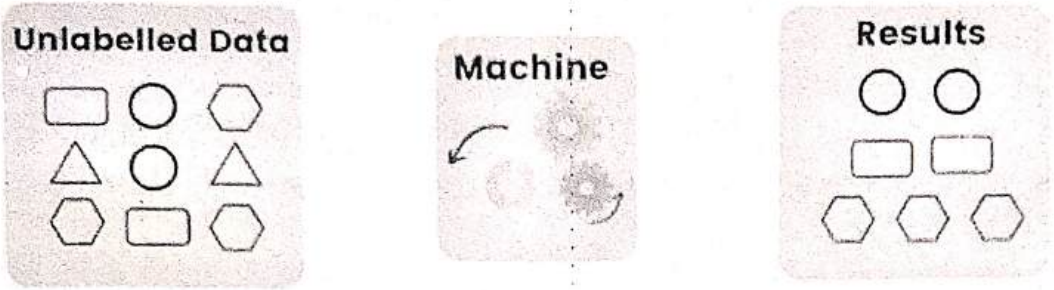

Fig 3: Un-Supervised Learning (source: projectpro.io)

Unsupervised Machine Learning is classified into two types: Association and Clustering.
Association:A method called Association Rules Analysis can be used to determine how different elements are related to one another. This technique is a rule-based Machine Learning technique that finds out some very useful relations between parameters of a large data set. This technique is basically used for market basket analysis that helps to better understand the relationship between different products. Some common Association Rule Learning algorithms:

- **Apriori Algorithm:** Finds patterns by exploring frequent item combinations step-by-step.
- **FP-Growth Algorithm:** An Efficient Alternative to Apriori. It quickly identifies frequent patterns without generating candidate sets.
- **Eclat Algorithm:** Uses intersections of itemsets to efficiently find frequent patterns.
- **Efficient Tree-based Algorithms:** Scales to handle large datasets by organizing data in tree structures.

Clustering: In the clustering process, data points are grouped into clusters such that points within the same cluster are more similar to each other than to those in other clusters. Typically, objects are grouped based on their similarities and differences. Clustering divides data into groups where points in the same group are similar and different from those in other groups. Some common clustering algorithms:

- **K-means Clustering:** Groups data into K clusters based on how close the points are to each other.
- **Hierarchical Clustering:** Creates clusters by building a tree step-by-step, either merging or splitting groups.
- **Density-Based Clustering (DBSCAN):** Finds clusters in dense areas and treats scattered points as noise.

- **Mean-Shift Clustering:** Discovers clusters by moving points toward the most crowded areas.
- **Spectral Clustering:** Groups data by analyzing connections between points using graphs.

Dimension Reduction: Dimensionality reduction is the process of reducing the number of features in a dataset while preserving as much information as possible. This technique is useful for improving the performance of machine learning algorithms and for data visualization. some popular Dimensionality Reduction algorithms:

- **Principal Component Analysis (PCA):** Reduces dimensions by transforming data into uncorrelated principal components.
- **Linear Discriminant Analysis (LDA):** Reduces dimensions while maximizing class separability for classification tasks.
- **Non-negative Matrix Factorization (NMF):** Breaks data into non-negative parts to simplify representation.
- **Locally Linear Embedding (LLE):** Reduces dimensions while preserving the relationships between nearby points.
- **Isomap:** Captures global data structure by preserving distances along a manifold.

- **Semi-Supervised Learning:** Semi-supervised Machine Learning is a technique that serves as a 'middle ground' between supervised and unsupervised learning. It utilizes both labeled and unlabeled data during the training process, typically using a small amount of labeled data along with a large volume of unlabeled data. This approach enables the model to learn effectively and make accurate predictions on new, unseen data.

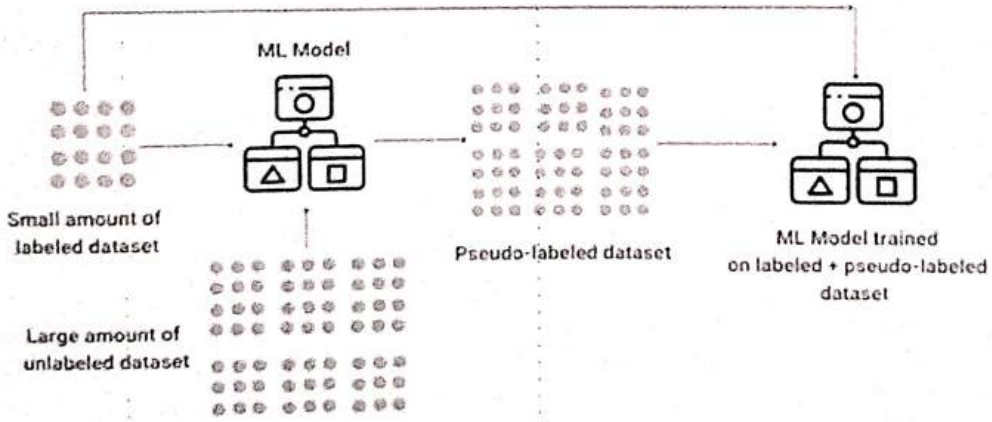
ML Model
Small amount of labeled dataset
Large amount of unlabeled dataset
Pseudo-labeled dataset
ML Model trained on labeled + pseudo-labeled dataset

Fig 4: Semi-Supervised Learning (source: projectpro.io)

- **Reinforcement Learning Algorithms:** Reinforcement Learning is another type of Machine Learning algorithm that closely resembles the way humans learn through trial and error. In Deep Reinforcement Learning, an agent interacts with its environment by performing actions and receiving feedback in the form of rewards or penalties. The key components of Reinforcement Learning are the Agent, the Environment, and the set of possible Actions.

- **The Agent:** The learner/decision-maker
- **The Environment:** Anything the agent interacts with
- **The Actions:** What the agent does

Reinforcement Learning algorithms do not rely on labeled data; instead, agents learn solely through interactions with their environment. Based on feedback, the agent takes various actions with the goal of maximizing cumulative positive rewards.

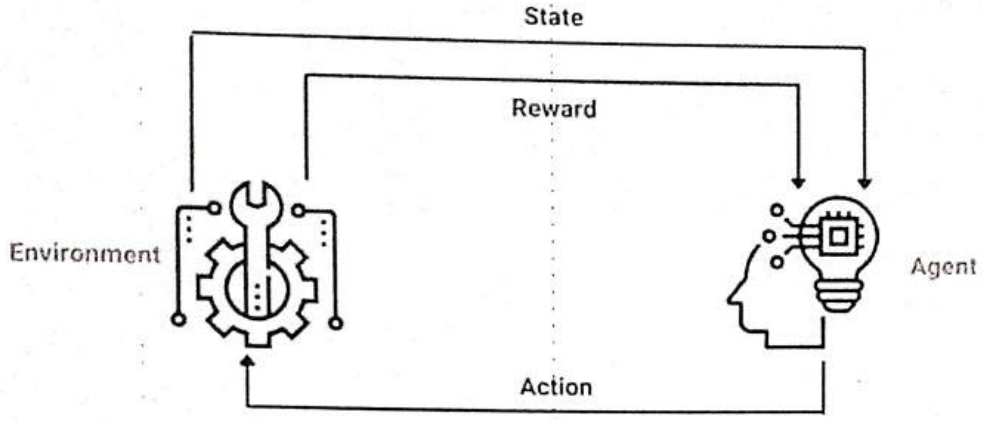

Fig 5: Reinforecement Learning (source: projectpro.io)

- **Machine Learning Model:**
  Data Set: Data about Learners is collected, such as Demographics Data: such as Age, gender, family income, parental education, etc. Academic Data: such as Previous grades, attendance, assignment scores, study time, participation in class. Behavioral Data: Library usage, Learning Activity, social interactions, Interest in learning etc.

  Data Preprocessing: prepares raw data for Machine Learning by cleaning, transforming, and structuring it to improve model accuracy and performance. Data Cleaning: Removing for filling missing data. Data Conversion: Convert text into numbers using Label Encoding. Normalization: Scale or normalize data so that all features have a similar range, improving model performance.

  Feature Extraction: turns raw data into useful features that help Machine Learning models work better by reducing complexity and focusing on important information. Feature Selection: Keep only important features to improve speed and reduce complexity. Data Splitting: Split data into training set and testing set to check how well the model performs on unseen data.

  Model Selection: in Machine Learning involves choosing the best model that performs well on both training and unseen data while avoiding overfitting or underfitting. Overfitting: A model that learns the training data too closely, including noise and random fluctuations, is said to be overfitting. Underfitting: A model that is too simple to capture the underlying patterns in the data is said to be underfitting.

  Training the Model: Training a model means teaching it to make accurate predictions by learning from labeled data and adjusting its parameters to reduce errors.

Prediction: In Machine Learning, prediction means using a trained model to guess or estimate an unknown output (target) based on new input data. The goal is to identify patterns and relationships in training data and apply them to make accurate forecasts or estimates about future data instances.

National-International Status

Globally, the prediction of Learners Academic Performance using Machine Learning is a growing research domain, especially in developed countries. US and Europe

- Top universities and EdTech companies use Machine Learning to find students who may struggle early and to give them personalized learning suggestions.
- Models like Decision Trees, SVM, Random Forest, and Deep Learning (e.g., LSTM (Long Short Term Memory), CNN) have been used on large educational datasets such as MOOCs (Massive Open Online Courses) to predict course completion and dropout risks. Increasing interest in Explainable AI (XAI) to make model decisions understandable for educators.
- Increasing interest in Explainable AI (XAI) to make model decisions understandable for educators.
- Countries (Asia-China, South Korea, Japan) emphasize student engagement and behavioral data collected via smart classrooms and Learning Management Systems (LMS) and Advanced deep learning models are used to forecast performance based on not only academic but also social and emotional factors.

Indian researchers are actively exploring Machine Learning to predict Learners performance, employability, and dropout rates in higher education. Datasets from UGC, AICTE-affiliated institutes, and government schemes (like SWAYAM) are being utilized. Study purpose uses Classification algorithms (Random Forest, Naive Bayes, Support Vector Machine) for Academic Performance and Regression models for predicting marks or CGPA.

Challenges in India

- Lack of data and privacy issues are big problems.
- Many colleges do not keep detailed or digitized data.
- A model trained in one region may not work well in another because of cultural, language, and academic differences.

Research Hypothesis

Primary Hypothesis (Main Hypothesis): H1: Artificial Intelligence models can accurately predict the Academic Performance of Learner's based on Historical, Academic and Behavioral data.

Supporting Hypotheses (Sub-Hypotheses):

- **H_1a:** Machine Learning algorithms such as Decision Tree, Support Vector Machine, Random Forest, Neural Network provide higher accuracy in predicting Learner's performance compared to traditional statistical methods.
- **H_1b:** Incorporating non-academic factors such as attendance, socio-economic background, and learning behavior improves the prediction accuracy of Artificial Intelligence models.
- **H_1c:** Early prediction of academic risk using Artificial Intelligence allows timely interventions that improve the final academic outcomes of Learner's.
- **H_1d:** Feature selection techniques significantly impact the performance of Artificial Intelligence models in predicting Learner's grades.

Components of Hypothesis

- **Use of Artificial Intelligence (AI)** techniques, such as Machine Learning (ML) algorithms, Data Analytics Tools, or Predictive Models.
- **Academic Performance** of Learner's (e.g., GPA, marks, pass/fail status, classification like Distinction/First Class).
- **Level of students** in a specific educational institution, region, or field (e.g., engineering, computer science).

Research Methodology

To predict a Learner's Academic success or failure (such as final grades, pass/fail status, dropout risk) based on Historical and Demographic data.

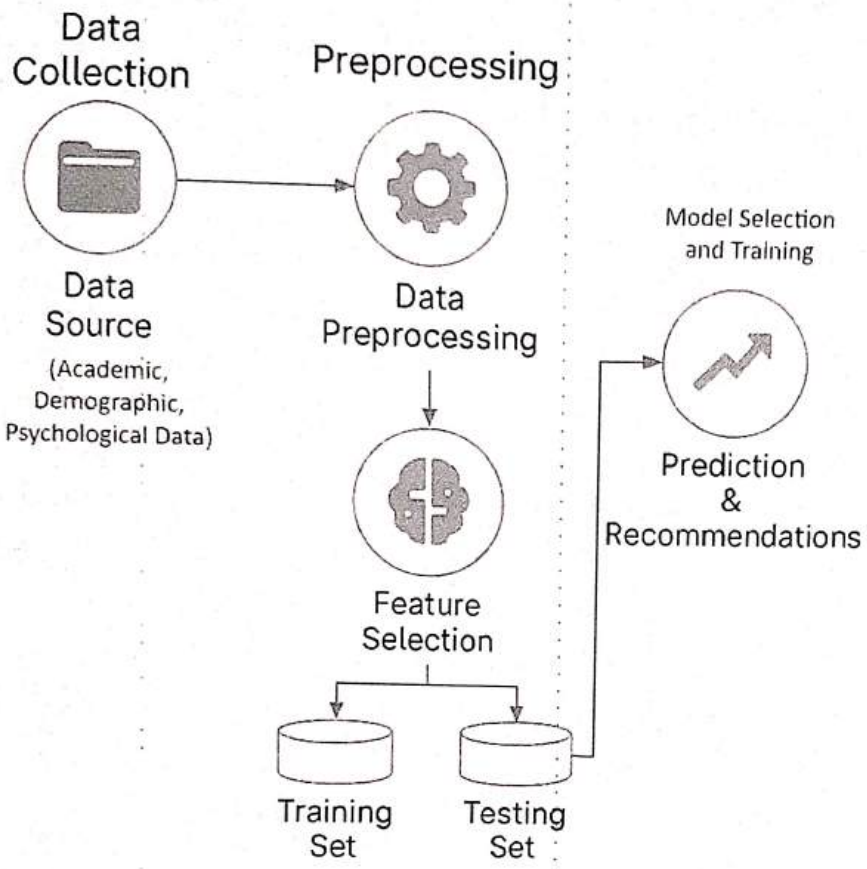

Fig 6: Machine Learning Model

1. **Data Gathering:** Data about Learner's is collected, such as:

   - **Academic Data:** Collect Learner's academic records such as marks, Previous grades, attendance, assignment scores, study time, participation in class.
   - **Demographics Data:** such as Age, gender, family income, parental education, etc.
   - **Behavioral Data:** Library usage, Learning Activity, social interactions, Interest in learning etc.

2. **Data Preprocessing:** Data Preprocessing in Machine Learning is the crucial first step of transforming raw data into a usable format for training Machine Learning models. It involves cleaning, transforming, and structuring data to address issues like missing values, inconsistencies, and irrelevant information. This process is essential for improving model accuracy, efficiency, and overall performance. Preprocessing include following:

   - **Data Cleaning:** Removing missing data, Duplicate data or inconsistent entries.
   - **Data Conversion:** Convert text into numbers using Label Encoding.

- **Normalization:** Scale or normalize data so that all features have a similar range, improving model performance.

3. Feature Extraction: Feature Extraction in Machine Learning is the process of selecting and transforming raw data into a set of meaningful features that can be used by machine learning models. It's a crucial step to improve model performance by simplifying the data, reducing dimensionality, and highlighting the most relevant information.

- **Feature Selection:** Keep only important features to improve speed and reduce complexity.
- **Data Splitting:** Split data into training set and testing set to check how well the model performs on unseen data.

4. Model Selection: Model selection in Machine Learning is the crucial process of choosing the best model from a set of potential candidates for a specific task and dataset. This involves evaluating different models based on their performance, complexity, and ability to generalize to new data. The goal is to find a model that not only performs well on the training data but also generalizes effectively to unseen data, avoiding issues like overfitting or underfitting. Different Machine Learning Algorithms can be used, such as:

Algorithm	Purpose/Reason
Decision Trees	Easy to understand, handles categorical and Numerical Data
Random Forest	Reduces overfitting, increases accuracy
Support Vector Machine (SVM)	Effective for complex boundaries
Logistic Regression	Simple, useful for pass/fail or yes/no prediction
Neural Networks (Deep Learning)	Captures complex patterns (for large datasets)

Table 1: Use of different ML Algorithm

Model Evaluation Matrix

- Accuracy
- Precision Recall
- F1-Score
- Confusion Matrix

5. Training the model: Training a model in Machine Learning involves teaching a Machine Learning algorithm to make predictions or decisions by learning from data. This process involves feeding labeled data to the algorithm, which then iteratively adjusts its internal parameters to minimize the difference between its predictions and the actual outcomes.

- Split dataset into Training and Testing sets (e.g., 80/20)
- Feed training data to the selected algorithm
- Use cross-validation to avoid overfitting

6. Prediction: In Machine Learning, prediction means using a trained model to guess or estimate an unknown output (target) based on new input data. The goal is to identify patterns and relationships in training data and apply them to make accurate forecasts or estimates about future data instances. Evaluation with metrics like accuracy, precision, recall, F1-score

Statistical Plan

Objective: To Develop a Machine Learning model that accurately predicts academic performance of Learner's using relevant Academic, Behavioral, and Demographic Data.

1. Study Design:

- Type: Quantitative, Observational, Retrospective study
- Approach: Predictive analytics using Supervised Machine Learning Techniques.
- Unit of Analysis: Individual Learner's records.

2. Data Collection Plan:

- **Institutional Academic Data:** Colleges managing the data that are deploying by us, such records like Attendance, Grades, Assignment scores, Internal/External Exam scores, Demographic data (age, gender, socioeconomic status) and Learner's Behaviours.
- **Surveys:** Fetching the data based on factors such as study habits, family background, and personal interests.

3. Data Preprocessing

- **Data Cleaning:** Through this, we hold the missing values and errors in data.
- **Encoding:** The conversion of categorical variables into numerical format with the help of one-hot encoding or label encoding (eg. Gender, course name).
- **Feature Scaling:** Standardization or min/max normalization numerical attributes for the definite algorithms like SVM (Support Vector Machine) or KNN (K-Nearest Neighbour).

4. Feature Selection: We detect the most common variable which impacts on performance of student. The possible features involves like:

- Attendance rate
- Past Academic Performance
- Participation in extracurricular activities
- Study hours per week
- Parental Education and Occupation
- Allowance for learning resources such as private tutoring and books.
- Chi-square Test: For categorical features

5. Model Selection and Development: The models are classified into two, they are:
Regression Models: This model is used when we are supposed to predict the constant output like final grade. Some techniques involves in this area:

- Linear Regression
- Decision Trees or Random Forests
- Gradient Boosting Machines

Classification Models: While we are predicting the categorical or absolute outcomes like "Pass/Fail". The tool are:

- Logistic Regression
- SVM (Support Vector Machine)
- K-Nearest Neighbours (KNN)
- Neural Networks (optional, depending on data size)

6. Training the Model

- Dataset is separated into the sets of Training and Testing sets. (Train-Test Split: 70:30 or 80:20)
- Our model is trained on the training set and validating or checking the performance.

7. Model Evaluation

- Accuracy
- Precision, Recall, F1-score
- Confusion Matrix
- ROC-AUC Curve (for binary classification)
- The regression task is performing by us with few metrics like Mean Absolute Error (MAE) or Root Mean Square Error (RMSE-for regression based GPA prediction).

8. Statistical Tests: For insights and comparisons:

- ANOVA / T-Test: To compare mean differences in performance groups
- Chi-square Test: To examine associations between categorical features and performance
- Correlation Matrix: For multicollinearity detection

9. Software/Tools Used

- Python (Scikit-learn, Pandas, NumPy, Seaborn, XGBoost)

- OP. R (caret, randomForest, ggplot2)
- Visualization: Matplotlib, Tableau/PowerBI (for final report)

10. Result Interpretation

- Identify most significant predictors (e.g., attendance, prior marks)
- Compare model performances statistically
- Provide recommendations for academic interventions

11. Deployment and Feedback

- By integrating the prediction model, the feedback is must gathered from the educators and Learner's about our model's performance based on predictions.
- Retrain our created model usually with fresh data and make sure of its relevancy and accuracy.

12. Report/Documentation

- Dataset Summary
- Feature Dictionary
- Code and model versioning

Limitations of Study

- **Incomplete Data:** Missing or inaccurate Academic, Demographic, or Behavioral Data can mislead the model.
- **Privacy Concerns:** Collecting sensitive Learner's (health, mental status, personal issues) is ethically and legally constrained.
- **Institution-Specific Models:** ML models trained on one Institution's data may not perform well when applied elsewhere due to curriculum, assessment, or Learner's diversity differences.
- **Changing Policies and Curricula:** Education systems evolve, and Machine Learning model may become outdated quickly if not updated to reflect these changes.
- **Unpredictable Human Behavior:** Learner's choices, motivation, and life circumstances are complex and can change unexpectedly, which models struggle to capture.
- **Lack of Actionable Insights:** Even if a prediction is accurate, Machine Learning models may not suggest clear interventions to improve Learner's outcomes.
- **Need for Continuous Monitoring:** Model performance degrades over time if not regularly updated with fresh data.

Expected Results/Outcomes

The proposed framework focus on analyzing and merging the Learenr's Demographics Data (place of residence, hobby, family size, etc.), Personal Data (admission details, age, gender, address, academic results, etc.). Academic Data, and Psychological Data (ability, attitude, behavior, motivation, etc.) to extract all possible knowledge from Learner's, teachers and parents. After surveying and selecting dataset, we predict the student academic performance and accuracy using Machine Learning and Deep Learning.

- Accurate Prediction Models: Machine Learning models like Random Forest, Support Vector Machine, and Neural Networks will predict Learner's performance with over 90% accuracy.
- Key Predictive Factors Identified: Important features such as attendance and internal marks will be highlighted through feature importance analysis.
- Algorithm Comparison: Performance metrics (Accuracy, Precision, Recall) will help determine the best-performing model.
- Improved Learner's Support: Insights will aid in designing personalized interventions like mentoring or Remedial Classes.
- Customizable Models: Models can be adapted for specific subjects, semesters, or Learner's groups.
- Enhanced Academic Planning: Institutions can use predictions for curriculum design and resource planning.
- Research Contribution: Potential for publication or presentation in academic forums related to Artificial Intelligence in education.

References

[1] Jitender Kumar, Ritu Vashistha, Kushwant Kaur, Siroj Kumar Singh,"Machine Learning Techniques of Predicting Student's Performance", 2023 International Conference in Advances in Power, Signal, and Information Technology (APSIT) — 979-8-3503-3936-9/23/ ©2023 IEEE — DOI: 10.1109/APSIT58554.2023.10201696

[2] Sathiyapriya N, Rithika C, Sanjay S, Subathra S, "Utilizing Machine Learning to forecast a students performance", 2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA) — 979-8-3503-4922-1/24/ ©2024 IEEE — DOI: 10.1109/AIMLA59606.2024.10531452

[3] Mohd Fazil, Ang elica R ısquez, Claire Halpin, "A Novel Deep Learning Model for Student Performance Prediction Using Engagement Data", Journal of Learning analytics, Volume 11(2), 23–41, https://doi.org/10.18608/jla.2024.7985, ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License

[4] Kuburat Oyeranti Adefemi, Murimo Bethel Mutanga and Vikash Jugoo, "Hybrid Deep Learning Models for Predicting Student Academic Performance", Mathematical and Computational Applications Appl. 2025, 30, 59. https:// doi.org/10.3390/mca30030059

[5] John-Bosco Diekuu, M. S. Mekala, Ulric Sena Abonie, John Isaacs and Eyad Elyan, "Predicting student next-term performance in degree programs using AI-based approach: a case study from Ghana", COGENT EDUCATION 2025, VOL. 12, NO. 1, 2481000, https://doi.org/10.1080/2331186X.2025.2481000

[6] Ujjwal Biswas, Sweta Sinha, Piyali De, "Recurrent Neural Network Model to Predict and Monitor Student Performance", 2025 8th International Conference on Electronics, Materials Engineering and Nano-Technology (IEMENTech) — 979-8-3315-2076-2/25/ ©2025 IEEE — DOI: 10.1109/IEMENTECH65115.2025.10959550

[7] Revathy P, Priya M, "Predictive and Prescriptive Analytics in Education: A Deep Dive into Machine Learning and Deep Learning Applications ", 2025 7th International Conference on Signal Processing, Computing and Control (ISPCC) — 979-8-3315-3893-4/25/ ©2025 IEEE — DOI: 10.1109/ISPCC66872.2025.11039382

[8] Mohammad AlShaikh-Hasan, Gheorghita Ghinea, "Predicting Student Performance Using Deep Learning and Machine Learning Techniques", 2025 1st International Conference on Computational Intelligence Approaches and Applications (ICCIAA) — 979-8-3315-2365-7/25/ ©2025 IEEE — DOI: 10.1109/ICCIAA65327.2025.11013240

[9] Neeta Sharma, M. K. Sharma, Umang Garg, "Predicting Academic Performance of Students Using Machine Learning Models", 2023 International Conference on Artificial Intelligence and Smart Communication (AISC) — 979-8-3503-2230-9/23/ ©2023 IEEE — DOI: 10.1109/AISC56616.2023.10085214

[10] Naveed Anwer Butt Mejdl Safran ,Zafar Mahmood, Khawar Shakeel, Sultan Alfarhood , Andimranashraf, "Performance Prediction of Students in Higher Education Using Multi-Model Ensemble Approach", IEEE Access, Digital Object Identifier 10.1109/ACCESS.2023.3336987

[11] George Bennett Winarto, Riky Supriyadi, Alexander Agung Santoso Gunawan, Muhammad Edo Syahputra, " Machine Learning-Driven Predictions to Enhance Academic Performance in High School Students: Systematic Literature Review", 2024 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS) — 979-8-3315-3324-3/24/©2024 IEEE — DOI: 10.1109/ICIMCIS63449.2024.10957540

[12] Nihel Fatima Baarir, Samir Bourekkache, Ahmed Aloui, " A Boosting Machine Learning Model for Predicting Student Performance in E-Learning", 2025 International Symposium on iNnovative Informatics of Biskra (ISNIB) — 979-8-3315-0957-6/25/ ©2025 IEEE — DOI: 10.1109/ISNIB64820.2025.10983182

[13] Delali Kwasi Dake, Esther Gyimah, Godwin Kudjo Bada, "A Deep Learning Model to Predict Information Technology Students Learning Style in Adaptive Learning Systems", 2025 4th International Conference on Sentiment Analysis and Deep Learning (ICSADL) — 979-8-3315-2392-3/25/ ©2025 IEEE — DOI: 10.1109/ICSADL65848.2025.10933094

[14] Sheng Qu, "Behavior Prediction and Management Algorithm of Students in Technical Colleges Based on Deep Reinforcement Learning", 2024 International Conference on Telecommunications and Power Electronics (TELEPE) — 979-8-3503-6921-2/24/ ©2024 IEEE — DOI: 10.1109/TELEPE64216.2024.00032

[15] V T Ram Pavan Kumar M, N. Giridhar, S. Pavani, R. Srilekha, S. Harika, P. Meghana,"A Machine Learning Framework for Evaluating College Performance and Student Outcomes", 2025 8th International Conference on Trends in Electronics and Informatics (ICOEI) — 979-8-3315-4460-7/25/ ©2025 IEEE — DOI: 10.1109/ICOEI65986.2025.11013685

[16] Fahim Uz Zaman, Muhammad Zohaib Khan, Aisha Imroz, Abdullah Ayub Khan, Muqaddas Salahuddin Maria Kainat, Maria Kainat," Student Performance Analysis in Higher Education Using Integrated Approach of Machine Learning Techniques", 2024 International Conference on Sustainable Technology and Engineering (i-COSTE) — 979-8-3315-1733-5/24/ ©2024 IEEE — DOI: 10.1109/I-COSTE63786.2024.11024835

[17] Shilpa Mangesh Pande, "Machine Learning Models for Student Performance Prediction", 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA) — 979-8-3503-9720-8/23/ ©2023 IEEE — DOI: 10.1109/ICIDCA56705.2023.10099503

[18] Anand Tripathi, "Development of an Intelligent System for Predictive Analysis of Student Academic Performance using Machine Learning Techniques", 2025 International Conference on Inventive Computation Technologies (ICICT) — 979-8-3315-1224-8/25/ ©2025 IEEE — DOI: 10.1109/ICICT64420.2025.11005033

[19] Gonuguntla Lahari, Rajashree, Santhanakrishnan, Jemshia Miriam, Nafees Muneera, "Integrative Machine Learning System for Comprehensive Student Personality Assessment and Academic Performance Enhancement", 2025 6th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI) — 979-8-3315-2266-7/25/ ©2025 IEEE — DOI: 10.1109/ICMCSI64620.2025.10883513

[20] Sreeramaneni Chandrahasa, Sonti Meghala Sri Ganeshan, Rahul Chowdary Maddineni, Mellampudi Sai Divya, Praveen Tumuluru, Bulla.Suneetha, "Machine Learning Algorithms based Student

Performance Prediction based on Previous Records", 2023 7th International Conference on Computing Methodologies and Communication (ICCMC) — 978-1-6654-6408-6/23/ ©2023 IEEE — DOI: 10.1109/ICCMC56507.2023.10084099

[21] Gheed M. Alsalem, Noor Sarhan, Mustafa Hammad, Bayan Zawaideh, "Predicting Students' Performance using Machine Learning Classifiers", 2024 25th International Arab Conference on Information Technology (ACIT) — 979-8-3315-4001-2/24/ ©2024 IEEE — DOI: 10.1109/ACIT62805.2024.10877102

[22] Yu Liu, Yanchuan Hui, Dongxu Hou, Xiao Liu, "A Novel Student Achievement Prediction Method Based on Deep Learning and Attention Mechanism", January 2023IEEE Access PP(99):1-1 DOI:10.1109/ACCESS.2023.3305248

[23] Jialun Pan, Zhanzhan Zhao, and Dongkun Han, "Academic Performance Prediction Using Machine Learning Approaches: A Survey", January 2025IEEE Transactions on Learning Technologies PP(99):1-18 DOI:10.1109/TLT.2025.3554174

[24] Essa Alhazmi And Abdullah Sheneamer, "Early Predicting of Students Performance in Higher Education" January 2023IEEE Access PP(99):1-1 DOI:10.1109/ACCESS.2023.3250702

[25] Mai Abdalkareem and Nasromin-Allah, "Explainable Models for Predicting Academic Pathways for High School Students in Saudi Arabia", February 2024, VOLUME 12, 2024 , Digital Object Identifier 10.1109/ACCESS.2024.3369586

[26] Lidya R. Pelima, Yuda Sukmana, Yusep Rosmansyah, "Predicting University Student Graduation Using Academic Performance and Machine Learning: A Systematic Literature Review", January 2024IEEE Access PP(99):1-1 DOI:10.1109/ACCESS.2024.3361479